Genetic Linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* Using a Pseudo-Testcross: Mapping Strategy and RAPD Markers

Dario Grattapaglia and Ronald Sederoff

Forest Biotechnology Group and Departments of Genetics and Forestry, Box 8008, North Carolina State University, Raleigh, North Carolina 27695.

Manuscript received January 10, 1994
Accepted for publication May 3, 1994

ABSTRACT

We have used a "two-way pseudo-testcross" mapping strategy in combination with the random amplified polymorhic DNA (RAPD) assay to construct two moderate density genetic linkage maps for species of *Eucalyptus*. In the cross between two heterozygous individuals many single-dose RAPD markers will be heterozygous in one parent, null in the other and therefore segregate 1:1 in their F_1 progeny following a testcross configuration. Meiosis and gametic segregation in each individual can be directly and efficiently analyzed using RAPD markers. We screened 305 primers of arbitrary sequence, and selected 151 to amplify a total of 558 markers. These markers were grouped at LOD 5.0, $\theta = 0.25$, resulting in the maternal *Eucalyptus grandis* map having a total of 240 markers into 14 linkage groups (1552 cM) and the paternal *Eucalyptus urophylla* map with 251 markers in 11 linkage groups (1101 cM) ($n = 11$ in Eucalyptus). Framework maps ordered with a likelihood support ≥1000:1 were assembled covering 95% of the estimated genome size in both individuals. Characterization of genome complexity of a sample of 48 mapped random amplified polymorphic DNA (RAPD) markers indicate that 53% amplify from low copy regions. These are the first reported high coverage linkage maps for any species of *Eucalyptus* and among the first for any hardwood tree species. We propose the combined use of RAPD markers and the pseudo-testcross configuration as a general strategy for the construction of single individual genetic linkage maps in outbred forest trees as well as in any highly heterozygous sexually reproducing living organism. A survey of the occurrence of RAPD markers in different individuals suggests that the pseudo-testcross/RAPD mapping strategy should also be efficient at the intraspecific level and increasingly so with crosses of genetically divergent individuals. The ability to quickly construct single-tree genetic linkage maps in any forest species opens the way for a shift from the paradigm of a species index map to the heterodox proposal of constructing several maps for individual trees of a population, therefore mitigating the problem of linkage equilibrium between marker and trait loci for the application of marker assisted strategies in tree breeding.

THE genus *Eucalyptus*, family Myrtaceae, is native to Australia and adjacent islands. It includes over 700 species of hardwood trees and shrubs (BOLAND *et al.* 1984) of which approximately 30 are grown throughout the world for fiber and energy production. Due to their fast growth and wide range of adaptability, eucalypts are the most widely used species for plantation establishment in tropical and subtropical regions of the world. Today, they constitute the majority of the world's exotic hardwood forest and one of the world's main sources of biomass (ELDRIDGE *et al.* 1993).

In spite of its commercial importance and a worldwide effort in breeding and propagation research, very little effort has been devoted to the development and use of molecular genetic markers for species of this genus. Isozyme markers have been used for the study of mating systems in natural and exotic populations of Eucalyptus (MORAN and BELL 1983; FRIPP *et al.* 1987; SAMPSON *et al.* 1990; YEH *et al.* 1983). Isozymes are inexpensive and technically accessible markers, however they are limited when a broader genome coverage is required.

In the last decade, detailed linkage maps based on restriction fragment length polymorphism (RFLP) markers have been developed for a number of species [see reviews by STUBER (1992) and GRATTAPAGLIA (1994)]. To date only two studies analyzed RFLP patterns in Eucalyptus: a phylogenetic analysis of chloroplast DNA (STEANE *et al.* 1991), and the development and screening of a genomic library for RFLP probes (WOLFF *et al.* 1993). RFLP analysis requires the "up front" development of probe libraries and involves Southern blot hybridization, procedures that are time consuming and labor intensive. The advent of a polymerase chain reaction-based arbitrarily primed genetic assay called RAPD, arbitrarily primed polymerase chain reaction (AP-PCR) or DNA amplification fingerprinting (DAF) (WILLIAMS *et al.* 1990; WELSH and MCCLELLAND 1990; CAETANO-ANOLLÉS *et al.* 1991), has provided an extremely efficient way to detect DNA polymorphisms and generate large numbers of molecular markers for genetic mapping and genomic fingerprinting applications (TINGEY and DELTUFO 1992). RAPD markers do not require prior sequence information and can be viewed

directly by agarose gel electrophoresis without the need of specific probe libraries and radioisotope detection. RAPD markers are therefore much more accessible to the nonmolecular biologist than RFLPs.

In recent years, RAPD markers have allowed a significant advance in the ability to generate linkage maps quickly. Maps of RAPD markers were reported for loblolly pine (GRATTAPAGLIA *et al.* 1991), white spruce (TULSIERAM *et al.* 1992), Arabidopsis (REITER *et al.* 1992), apple (LAWSON *et al.* 1992) sugar cane (AL-JANABI *et al.* 1993), faba bean (TORRES *et al.* 1993), slash pine (NELSON *et al.* 1993) and peach (CHAPARRO *et al.* 1994). In Eucalyptus, RAPD markers have been used in genetic analyses of individuals and populations including clone fingerprinting, outcrossing rate estimation and phylogenetic relationship studies (GRATTAPAGLIA *et al.* 1992a). In Eucalyptus, no reports of genetic linkage of morphological or isozyme markers are available. However, in the last 2 years the construction of linkage maps of molecular markers has been undertaken for *Eucalyptus grandis* and *Eucalyptus urophylla* (GRATTAPAGLIA and SEDEROFF 1992), *Eucalyptus globulus* (SONG and CULLIS 1992) and *Eucalyptus nitens* (MORAN *et al.* 1992).

Most linkage maps in plants have been obtained from segregating populations derived from crosses between inbred lines. Such populations or even three-generation pedigrees are generally not available in trees and are difficult to obtain due to a significant genetic load and time constraints. In spite of this obstacle, existing three-generation outbred *Pinus taeda* and inbred Populus pedigrees (DEVEY *et al.* 1991; BRADSHAW and STETTLER 1993) have been used as mapping populations. Alternatively in conifers, haploid megagametophytes have allowed the direct analysis of linkage in gametes and the construction of genetic maps (CONKLE 1981; GRATTAPAGLIA *et al.* 1991; TULSIERAM *et al.* 1992; NELSON *et al.* 1993). Available pedigrees for the majority of angiosperm outbred tree species, including the eucalypts, generally involve only two parents and their full-sibs, or maternal half-sib families. To incorporate molecular marker assisted strategies into forest tree breeding it is imperative to explore alternative approaches for the construction of linkage maps that make use of pedigree structures already existing and commonly generated in tree breeding programs. Particularly in the case of trees, genetic linkage maps could prove a very powerful tool for accelerating breeding through marker assisted selection and recombination. Breeding programs of *Eucalyptus* species are based on simple or reciprocal recurrent selection, interspecific hybridization and, in some cases, clonal selection and deployment. Linkage disequilibrium generated by hybridization coupled with the possibility of capturing non additive genetic variance through clonal propagation are conditions that would greatly enhance the potential use of molecular marker assisted breeding strategies in these species.

Here we report the genetic inheritance, segregation and linkage of 558 single-dose RAPD markers for two closely related species of *Eucalyptus*, *E. grandis* Hill ex Maiden and *E. urophylla* S. T. Blake. We used a two-generation interspecific pedigree to construct single-tree genetic linkage maps using a two-way "pseudo-testcross" mapping strategy. We propose this mapping approach as a general strategy to generate single individual linkage maps quickly for highly heterozygous organisms. These are the first reported genetic linkage maps for Eucalyptus and among the first for any hardwood tree species.

MATERIALS AND METHODS

Plant material: A single controlled cross between two highly heterozygous elite trees was selected for genetic mapping. *E. grandis* (clone 44, Coffs Harbor provenance, Australia–selection from a Zimbabwe seed source), used as the female parent was crossed to *E. urophylla* (clone 28 selection from Rio Claro land race, Brazil), used as male, in 1989 at Aracruz Florestal S.A., Brazil. These two species belong to the same subgenus and section. They are easily crossed and their progeny are fully fertile. A second controlled cross used in the study was performed in the same year and location between a different female parent, *E. grandis* (clone 816/2 Atherton provenance, Australia) and the same *E. urophylla* male parent. Seeds obtained from these crosses were surface sterilized and germinated on solid agar containing half-strength MS medium (MURASHIGE and SKOOG 1962) under a 14-hr photoperiod. The mapping population consisted of 62 F_1 individuals. This population was immortalized by establishing clonal cultures of the individuals by vegetative propagation *in vitro* on half-strength MS medium supplemented with 0.005 mg/liter indol-butyric acid to stimulate rooting of microcuttings.

DNA extraction: Total genomic DNA was isolated from freeze dried adult leaf tissue of the parents and from fresh leaves of *in vitro* plantlet progeny using the protocol of DOYLE and DOYLE (1987), modified by the addition of 1% polyvinylpyrrolidone and 1% 2-mercaptoethanol to the isolation buffer. DNA concentration was estimated by gel electrophoresis comparing the fluorescence intensities of the ethidium bromide stained samples to those of λ DNA standards.

RAPD assay: Random 10-bp primers (kits OP-A through OP-Z) were obtained from Operon Technologies Inc. (Alameda, California). Amplification reactions (13 μl) were carried out according to WILLIAMS *et al.* (1990) with the following modifications: 0.4 μM 10-base primer, 10 μg/μl non-acetylated bovine serum albumin (New England Biolabs), 5–10 ng of genomic DNA and 1 unit of *Taq* DNA polymerase. With a genome size of 0.6 pg/1C (haploid genome DNA content) (GRATTAPAGLIA and BRADSHAW 1994) the amount of genomic DNA used in a RAPD reaction corresponded to between 8,000 and 16,000 haploid genome equivalents. Amplifications were performed in 96-well microtest plates using an MJ Research PT-100 thermal controller. RAPD products were analyzed by electrophoresis in 1.5% or 2.0% agarose gels containing 0.2 μg/ml ethidium bromide. On a custom made gel tray a full 96-well plate was run on a single gel. Gels were photographed under transmitted UV light using a MP4 Polaroid camera or an Eagleye™ video imaging system (Stratagene) and printed on 20 × 15-cm thermal paper. Gel scoring was performed directly from the photographs or thermal prints.

Primer screening: A total of 305 10-base random primers were screened against the two parents and a progeny sample

of six individuals. RAPD fragments that are polymorphic between the two parents and segregate in the progeny can be detected. With six individual progeny the probability of missing a polymorphic marker segregating 1:1 is 0.094. Twelve primers were conveniently screened on one 96-well plate and one gel. A total of 151 primers were selected during this step based on the number of RAPD polymorphisms amplified, their size and amplification intensity. Selected primers were used on the mapping population.

Scoring of RAPD markers: Segregation of RAPD markers in the mapping population was recorded in two independent replications, each one with a different set of individuals. In the first replicate the two parents and 30 progeny were assayed for RAPD markers with all the selected primers. In a second replicate, totally independent DNA extractions, reaction mixture preparations, gel analysis and genotype scoring were performed for another set of 32 progeny individuals. This procedure provided an internal control and an estimate of repeatability for all the scored markers. Markers that did not amplify consistently or could not be scored reliably across the two replicates were dropped from further analysis.

Segregating RAPD markers were identified by the manufacturer primer code corresponding to a particular 10-base sequence, followed by a number indicating the fragment size in base pairs. Fragment sizes were estimated using the software SEQAID II (RHOADS and ROUFA 1990). Following the fragment size, separated by a slash, a subjective score was given from 1 to 3 denoting the fragment amplification intensity, 3 being the most intense. For example RAPD marker A11_980/3 corresponds to a RAPD fragment amplified by Operon primer A11 (corresponding to the sequence 5′-CAATCGCCGT-3′), with size 980 bp, of high (score 3) amplification intensity. All the scored RAPD fragments were sampled from the agarose gel by gently stabbing the fluorescing band with a pipette tip and rinsing the tip into 20 μl of sterile TE buffer (10 mM Tris-HCl and 0.2 mM EDTA). All samples were stored at −20° until required for reamplification.

Characterization of genomic sequence complexity of RAPD marker loci: Total genomic DNA was extracted from a bulked leaf sample of individuals of the mapping population. In the wells of a dot-blot apparatus, an appropriate amount of DNA was denatured and vacuum transferred and cross-linked to a nylon membrane. Membranes were cut into strip blots with four contiguous dots containing 5, 0.5 and 0.05 μg of eucalypt DNA and 5 μg of herring sperm DNA as a negative control.

Dot blots to be used as controls for comparative analysis were prepared by performing a reconstruction experiment where known picogram amounts of pure RAPD fragment corresponding to 1, 10, 100 and 1000 copies were immobilized onto a membrane using 5.0 μg of herring sperm DNA as a carrier following the same contiguous dot blot format. Control blots were prepared for twelve different RAPD fragments varying in size from 300 to 1500 bp. A total of 48 RAPD fragments corresponding to mapped marker loci were labeled and used as probes on the dot blots to characterize the copy number of their internal sequences. Control blots were simultaneously probed with the corresponding fragment used in the reconstruction experiment. To further confirm the results, two RAPD fragments from each copy number class were then used as probes on genomic SOUTHERN (1975) blots.

RAPD fragment hybridizations: RAPD fragments to be used as hybridization probes were reamplified using the RAPD assay conditions described. Template DNA for reamplification consisted of a 3-μl volume of the 20-μl RAPD band sample. Nonradioactive labeling of the probes was performed with dUTP-digoxigenin according to manufacturer's recommendations (Boheringer Mannheim). Reamplified probes were checked for single band purity on a minigel. Probe hybridization, washes and detection procedures were carried out following manufacturer recommendations using the chemiluminescent substrate solution CSPD (Tropix Inc.).

Confirmation of RAPD marker inheritance, codominance and presence in different individuals: Putative codominance (size variation) of RAPD markers was investigated by DNA hybridization of gel blots of RAPD products with the putative codominant fragment used as probe. A subset of mapped RAPD markers were surveyed for their presence and segregation in a different individual of *E. grandis* by analyzing the second F_1 cross described previously. RAPD assay was carried out on the two parents and 10 progeny for each one of the two crosses. Confirmation of homology for RAPD markers of same size was carried out through DNA hybridization, using the fragment of interest as a probe. RAPD gels were blotted onto nylon membranes and hybridized as described.

Segregation and linkage analysis of RAPD markers: Segregating markers were scored for presence (1) or absence (2) of the amplified RAPD band. The parental origin of the marker was also recorded. Two separate data sets were obtained, one for each parent. In the pseudo-testcross configuration markers are present in one parent and absent in the other or vice versa, and are expected to segregate 1:1 in the F_1 generation. A χ^2 test ($\alpha = 0.05$) was performed to test the null hypothesis of 1:1 segregation on all scored markers. Preliminary grouping was done using a χ^2 test for independence of segregation at a threshold of 15.00 (~LOD = 5.0; R. DOERGE, personal communication). The linkage analysis was done using MAPMAKER (LANDER *et al.* 1987). The software program GMENDEL (LIU and KNAPP 1990) was also used during linkage analysis particularly for ordering linkage groups. To allow the detection of linkage of RAPD markers in repulsion phase the data set was duplicated and recoded. A LOD score of 5.0 and maximum $\theta = 0.25$ were set as linkage thresholds for grouping markers. Map distances in centimorgans were calculated using Kosambi's mapping function. Preliminary orders of marker loci in each linkage group were established using a matrix correlation method implemented by MAPMAKER. From this initial order, a subset of evenly spaced loci that could be ordered with a likelihood ratio support ≥1000:1 established a framework map. Error detection functions of MAPMAKER were employed to check potential genotyping errors in the framework markers. The final framework order obtained was then compared to the order outputted for the same subset of marker loci by GMENDEL that employs a simulated annealing algorithm. Markers that could not be ordered with equal confidence were indicated as accessory markers at an already specified locus on the map. Genome map sizes were estimated according to HULBERT *et al.* (1987) taking into consideration only pairwise comparisons between markers placed on the framework map.

RESULTS

Primer screening: In our standard conditions, RAPD reactions amplified an average of 10.7 visible bands on an ethidium bromide stained agarose gel. Primer screening was efficiently carried out using both parents and a sample of F_1 individuals (Figure 1). With this format, parental origin of the polymorphic loci as well as their allelic state (homozygous or heterozygous) was directly inferred from the presence of the fragment in one parent, the absence in the other and at least one presence/absence in the F_1 progeny sample. Of the 305 arbitrary primers screened, 57 (18.7%) did not yield any amplified product, 50 (16.4%) did not detect any visible

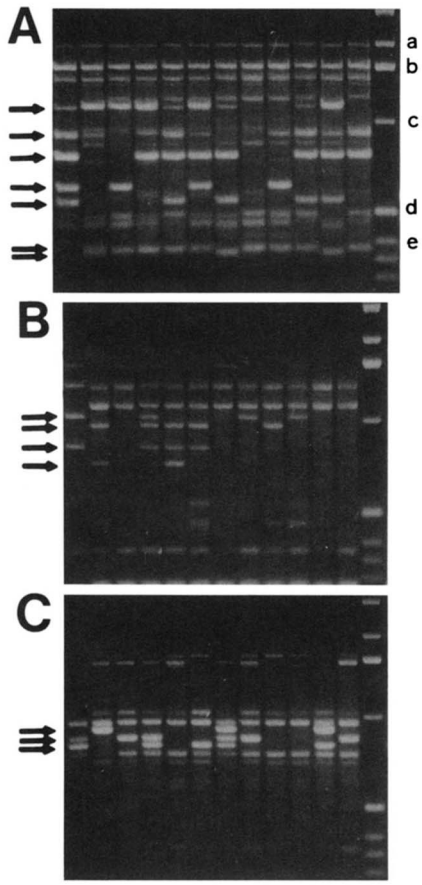


FIGURE 1.—Screening arbitrary 10-base primers (Operon Technologies Inc., Alameda, California) for pseudo-testcross marker configurations. Primers Y20 (A), Z11 (B) and R16 (C) were used in RAPD assays with genomic DNA of the parents *E. grandis* clone 44 and *E. urophylla* clone 28 (respectively, lanes 1 and 2 of each panel) and a progeny sample of 10 F_1 individuals. Segregating informative RAPD markers are indicated by arrows; last lane on the left is a 1-kb DNA ladder size standard (BRL) (a = 2036 bp; b = 1636 bp; c = 1018 bp; d = 506 bp; e = 396 bp).

sequence polymorphism in our particular cross and 198 (64.9%) uncovered at least one RAPD fragment polymorphism. From these 198 primers, a total of 151 were selected, aiming at maximizing the number of scorable markers per primer in the following mapping step. A total of 558 RAPD markers were scored on the mapping population, yielding an average of 3.69 markers/selected primer or 1.82 markers/any arbitrary primer. Similar ratios were recently reported for single-dose RAPD polymorphisms in sugar cane (1.88) (SOBRAL and HONEYCUTT 1993) and Stylosanthes (1.8) (KAZAN *et al.* 1993). Although the majority (64.9%) of primers screened detected at least one polymorphism, the screening step essentially doubled the time efficiency and halved the cost of data gathering in the mapping phase.

Scoring of markers on the mapping population: Segregation of a total of 558 RAPD markers was scored on the mapping population (Figure 2). The numbers of markers inherited from each parent were very similar: 272 from *E. grandis* and 286 from *E. urophylla*. RAPD fragments sizes ranged from 207 to 3335 bp, with an average of 979 ± 570 bp for *E. grandis* and 910 ± 521 bp for *E. urophylla*. The replicated design used (see MATERIALS AND METHODS) throughout the mapping phase yielded a repeatability estimate of 92.4%. From the 558 RAPD markers, 516 were fully repeatable in both replicates while 42 were not, 20 from *E. grandis* and 22 from *E. urophylla*. These markers either could be scored in the first set of progeny and not in the second or vice versa. With no exception, all of them were originally classified as class 1 markers denoting low amplification intensity and/or difficulty in scoring due to co-migrating fragments. Such fragments were not considered in further analysis.

Linkage analysis: Segregation ratios that departed from the Mendelian expectation of 1:1 at $\alpha = 0.05$ were detected at 10 marker loci in *E. grandis* and 10 marker loci in *E. urophylla*. No departure was detected at $\alpha = 0.01$. These apparently distorted markers (denoted by an asterisk following the marker identification) are clustered on only 2 linkage groups in *E. grandis* (groups 6 and 7) but are scattered in 6 linkage groups in *E. urophylla* (Figures 3 and 4). Note that at $\alpha = 0.05$, considering a total of 500 marker loci, around 25 of these are expected to display this behavior due to chance alone. Therefore, at this point we have no reason to suggest that such distortions have a biological basis. Such an indication may be drawn from the examination of a larger sample of meiosis. Furthermore, the parents of the mapped trees would be necessary to study the specific origin of the observed distortion, *i.e.*, excess of the allele of one grandparent *vs.* deficiency of the allele from the other. Only markers that passed the single-locus segregation test were initially used in the linkage grouping analysis. Distorted markers were later placed on the map by determining their most probable location in an already established gene order.

Linkage relationships of the segregating markers were established using both a χ^2 test for independence of segregation at a threshold value of 15.00 and by two-point mapping (LOD 5.0 and maximum $\theta = 0.25$). Both linkage analyses agreed very closely. Overall, linkages were robust at a LOD score range from 4.0 to 6.0. In view of the large number of markers, at lower LOD scores, especially below 4.0, occasional spurious linkages resulted in the agglomeration of some linkage groups, while increasing LOD scores beyond 6.00 would result in fragmentation of linkage groups. At LOD 5.0 the maternal *E. grandis* map has a total of 240 markers into 14 linkage groups and the paternal *E. urophylla* map 251 markers in 11 linkage groups (Figures 3 and 4). Twelve markers for *E. grandis* and 13 for *E. urophylla* remained unlinked at LOD 5.0. Although they were linked at a lower

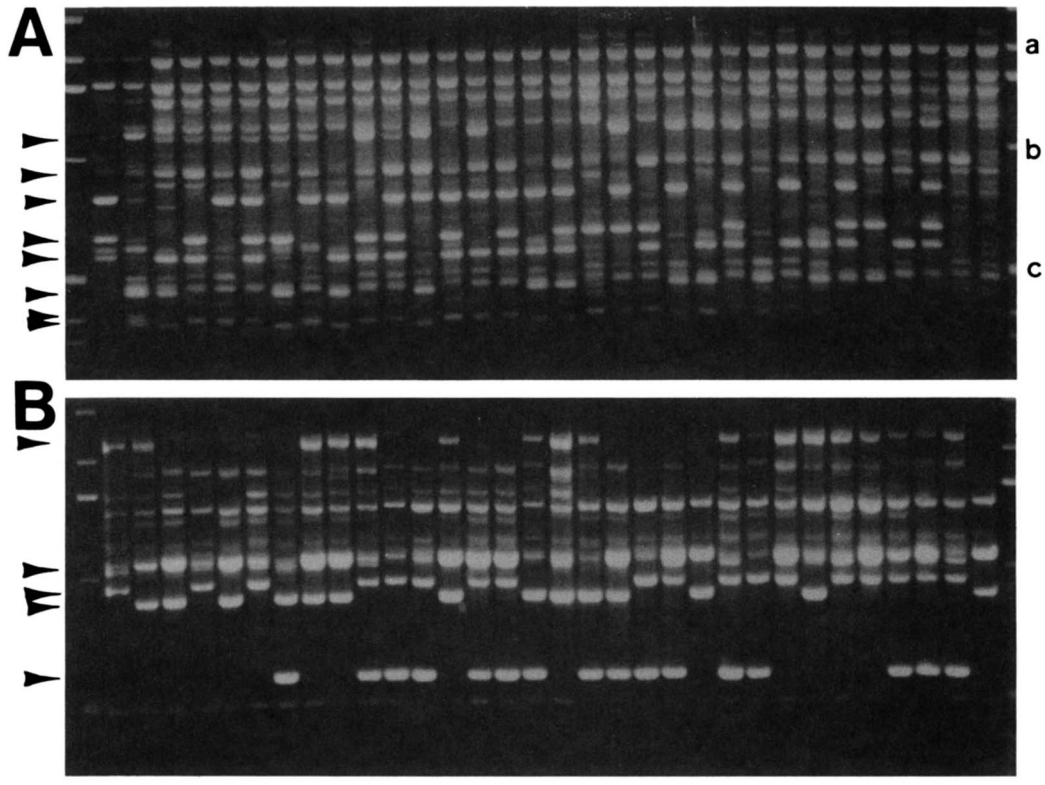
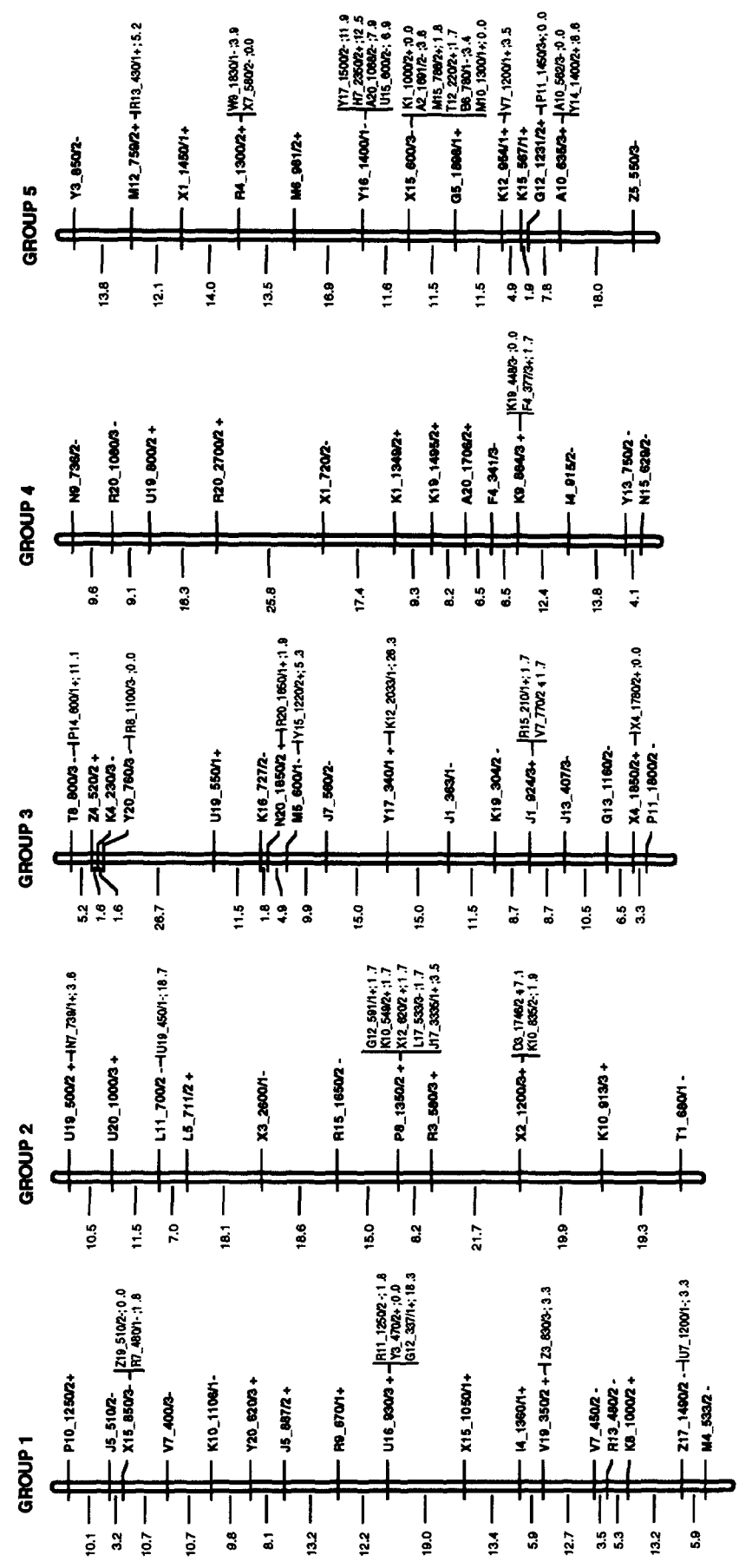


FIGURE 2.—Genetic inheritance and segregation of pseudo-testcross RAPD markers in the mapping population using a replication format (see MATERIALS AND METHODS). In both panels the first and last lanes are 1-kb DNA ladder size standards (BRL) (a = 2036 bp; b = 1018 bp; c = 506 bp). (A) RAPD assays on the parents (second and third lanes) and 30 F_1 progeny (first replicate format) using primer Y20; (B) 32 F_1 progeny (second replicate format) using primer A11. Segregating RAPD markers are indicated by arrows.

LOD (3.0), and higher θ (0.35), they did not contribute any significant additional information in terms of genome coverage and therefore they were left out of the final map versions. The proportion of unlinked markers found (4.5%) is smaller than those reported for other single-dose marker linkage mapping studies [*e.g.*, 10% in potato (GEBHARDT *et al.* 1989); 15.4% and 12.9% in sugar cane (AL-JANABI *et al.* 1993; DASILVA *et al.* 1993)].

Locus ordering and map construction: Both Eucalyptus species have $n = 11$ chromosomes, and therefore 11 linkage groups were expected in each map. This expectation was met for the *E. urophylla* map but not for the *E. grandis* map. However, for *E. grandis*, lowering the threshold LOD score to 4.0 and increasing θ to 0.35 would result in the merging of three pairs of linkage groups leading therefore to a correspondence between number of linkage groups and number of chromosomes. The following mergers with the respective highest LOD score between markers on separate groups would take place: groups 8 and 12 (LOD score 4.3 between markers N6_634/1 and K9_534/2); groups 11 and 13 (LOD score 4.4 between markers B6_759/1 and X12_530/2); groups 1 and 6 (LOD score 4.05 between V7_450/2 and L16_389/2) (Figure 3). In spite of the possibility of merging some linkage groups to attain the expected number of chromosomes, it was found more appropriate to assemble both maps with the same statistical stringency allowing for more meaningful comparisons between maps. An excess of linkage groups in relation to the haploid chromosome number has been reported for other species [*e.g.*, bean (NODARI *et al.* 1993); lettuce (KESSELI *et al.* 1990)].

The linkage groups were constructed using markers in both linkage phases. Markers on one linkage phase are indicated with a "+" sign following the marker identification code, while markers on the alternative phase are indicated with a "−" sign (Figures 3 and 4). Matrix correlation was used to get a preliminary locus ordering. Based on this first approximate order, a subset of candidate framework marker loci was selected spanning the whole linkage group at distances varying between 5 and 20 cM. This selection was based on a sequence of criteria, that by order of priority were as follows: (i) fragment intensity of amplification score; (ii) ease of marker scoring in view of co-migration of other fragments or smearing that could lead to errors in genotyping (gel photo was reviewed); (iii) number of missing data; (iv) size of amplified fragment (below 2000 and above 300 bp). Candidate framework markers selected were again ordered using matrix correlation. This linear order was then tested by permutating all possible sets of three


GROUP 1
P10_1250/2+
J5_510/2-
X15_850/3-
V7_400/3-
K10_1106/1-
Y20_620/3 +
J5_987/2 +
R9_670/1+
U16_930/3 +
X15_1050/1+
I4_1360/1+
V19_350/2 +
V7_450/2 -
R13_490/2 -
K9_1000/2 +
Z17_1490/2 -
M4_533/2 -
Z19_510/2-;0.0
R7_480/1-;1.8
R11_1250/2-;1.8
Y3_470/2+;0.0
G12_337/1+;18.3
GROUP 2
U19_500/2 +-IN7_738/1+;3.6
U20_1000/3 +
L11_700/2 -IU10_450/1-;18.7
L5_711/2 +
X3_2600/1-
R15_1850/2 -
P8_1350/2 +-
G12_591/1+;1.7
K10_549/2+;1.7
X12_920/2 +;1.7
L17_533/3-;1.7
J17_333/1+;3.5
R3_590/3 +
X2_1200/3+-
D3_1740/2 +7.1
K10_835/2-;1.9
K10_913/3 +
T1_680/1 -
GROUP 3
T8_800/3 -IP14_600/1+;11.1
Z4_520/2 +
K4_230/3 -
Y20_760/3 -IR8_1100/3-;0.0
U19_550/1+
K16_727/2-
N20_1850/2 +IR20_1650/1+;1.9
M5_600/1- IY15_1220/2+;5.3
J7_560/2-
Y17_340/1 +-IK12_2033/1+;26.3
J1_363/1-
K19_304/2 -
J1_924/3+-
R15_2101/1+;1.7
V7_770/2 +1.7
J13_407/3-
G13_1160/2-
X4_1850/2+ -IX4_1780/2+;0.0
P11_1800/2 -
GROUP 4
N9_736/2-
R20_1080/3 -
U18_800/2 +
R20_2700/2 +
X1_720/2-
K1_1349/2+
K19_1495/2+
A20_1706/2+
F4_341/3-
K9_864/3 +-
K19_449/3-;0.0
F4_377/3+;1.7
I4_915/2-
Y13_750/2 -
N15_629/2-
GROUP 5
Y3_850/2-
M12_759/2+ -IR13_430/1+;5.2
X1_1450/1+
R4_1300/2+-
W9_1830/1-;3.9
X7_580/2-;0.0
M8_961/2+
Y16_1400/1-
Y17_1500/2-;11.9
H7_2350/2+;12.5
A20_1080/2-;7.9
U15_600/2-; 6.9
X15_600/3-
K1_1000/2+;0.0
A2_1691/2-;3.6
M15_789/2+;1.8
T12_2202/+;1.7
G5_1898/1+
B6_7801/-;3.4
M10_1300/1+;0.0
K12_954/1+ -IV7_1200/1+;3.5
K15_567/1+
G12_1231/2+ -IP11_1450/3+;0.0
A10_638/3+ -IA10_583/3-;0.0
Y14_1400/2+;8.8
Z5_550/3-

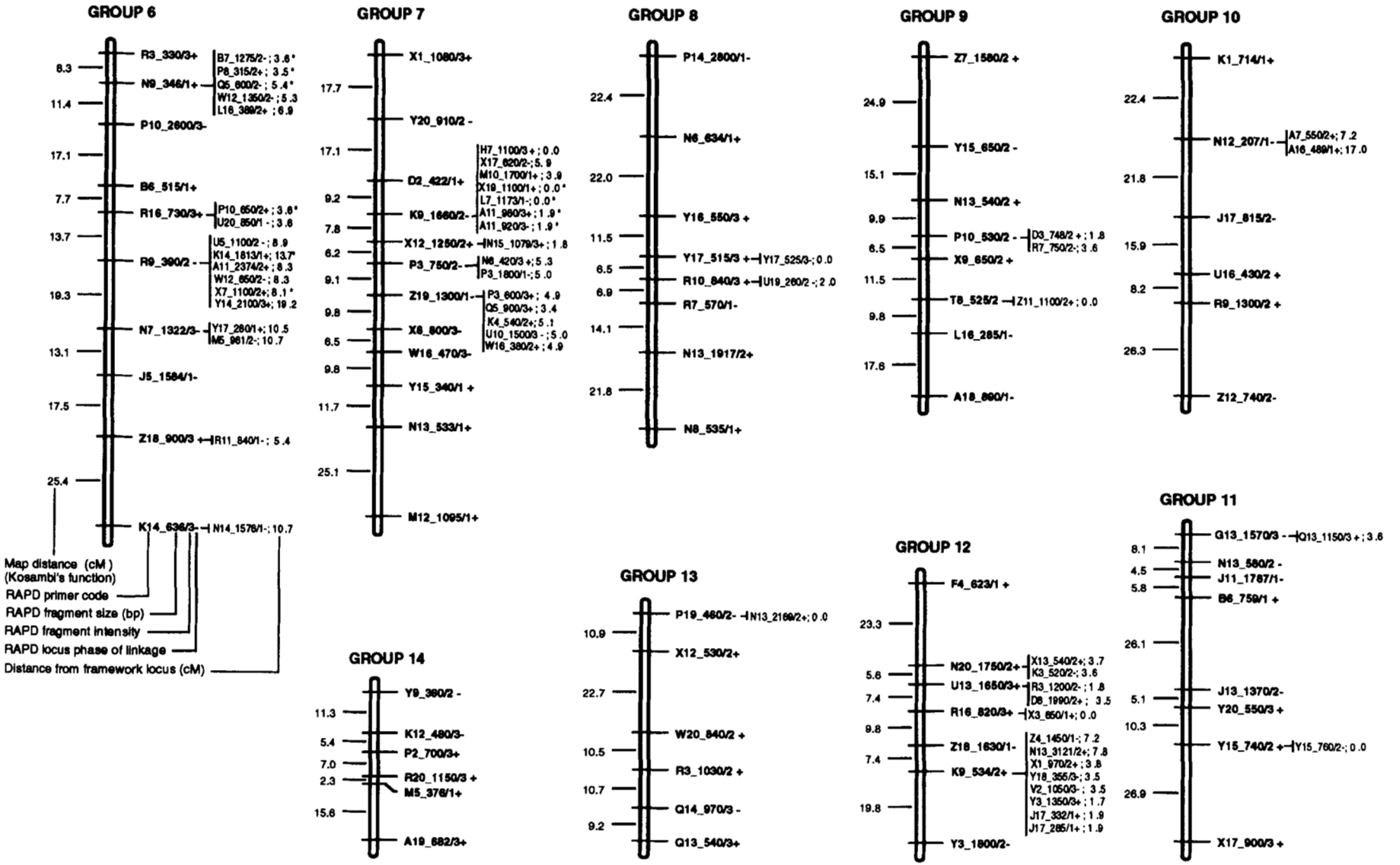
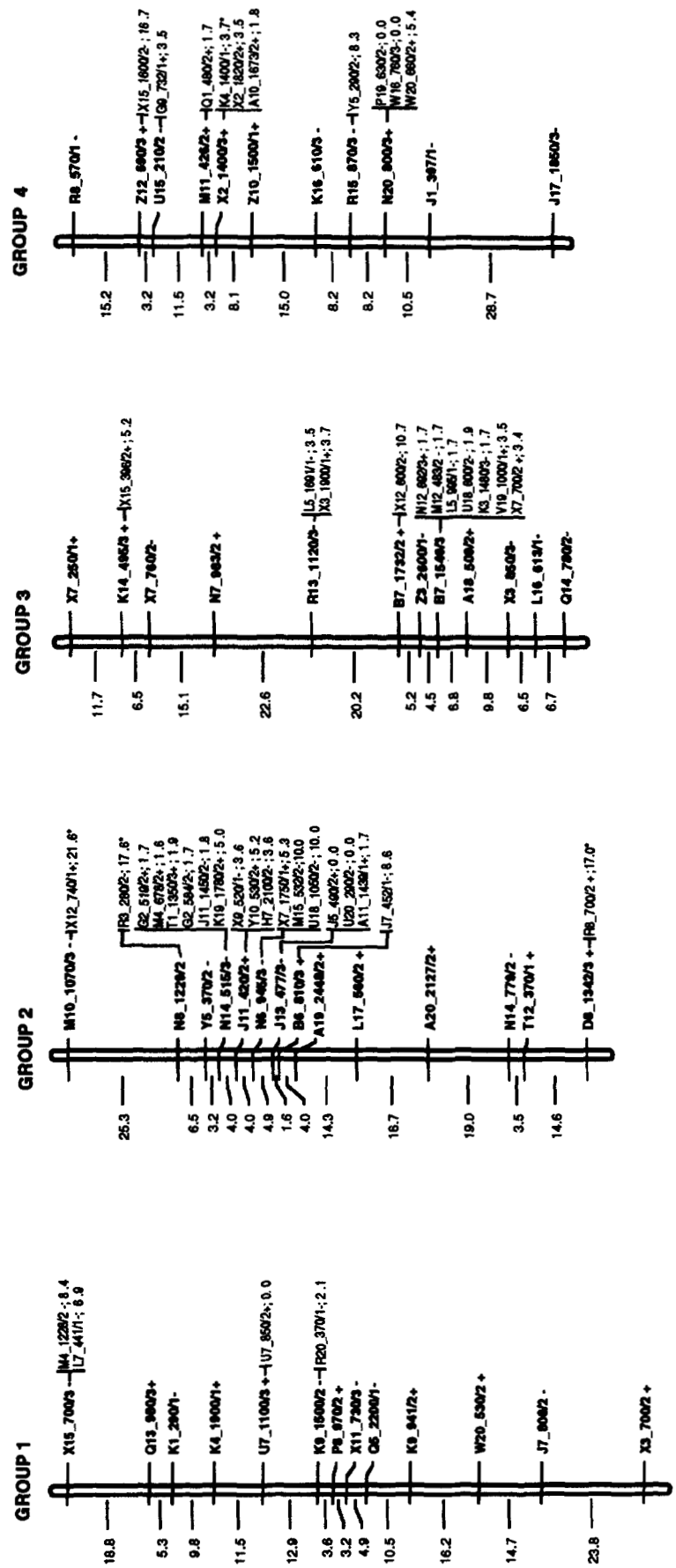


FIGURE 3.—Genetic linkage map of *E. grandis* clone 44. Linkage relationships of 240 RAPD markers in 14 linkage groups were established at a threshold LOD score 5.0 and maximum $\theta = 0.25$. A framework map of 142 loci or loci clusters (indicated in bold letters along the linkage groups) that could be ordered with a likelihood support $\geq$1000:1 was assembled covering 95.8% of the estimated genome size. RAPD marker loci that could not be ordered with equal confidence were designated as accessory markers and are listed on the right of the linkage groups along with the approximate cM distance to the closest framework locus. RAPD marker loci are identified by the operon primer code, fragment size in base pairs, fragment amplification intensity and linkage phase (+ or −) (see MATERIALS AND METHODS). A RAPD marker locus showing significant distortion from 1:1 segregation ratio, is indicated by an asterisk.


GROUP 1
GROUP 2
GROUP 3
GROUP 4

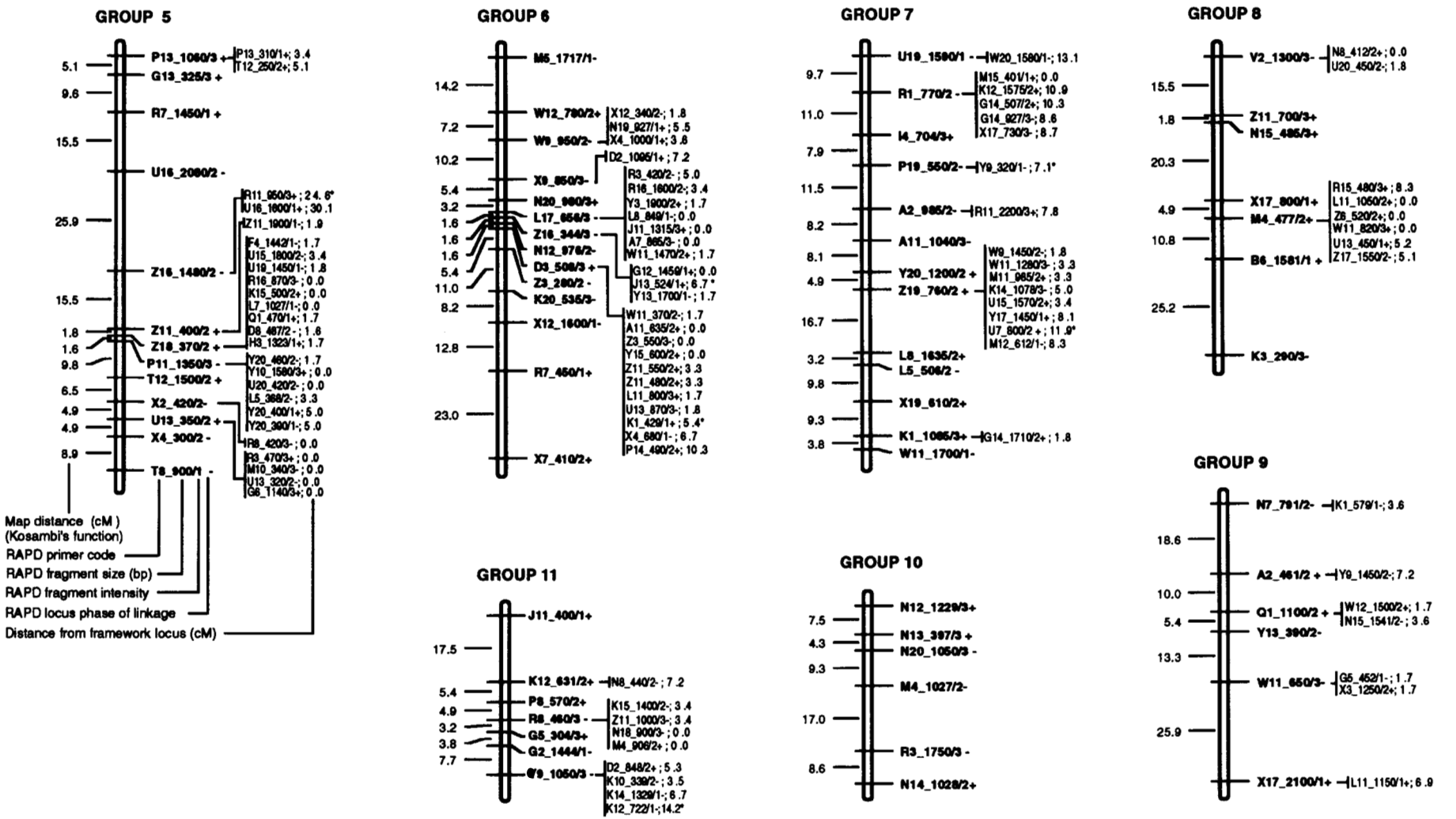


FIGURE 4.—Genetic linkage map of *E. urophylla* clone 28. Linkage relationships of 251 RAPD markers in 11 linkage groups were established at a threshold LOD score 5.0 and maximum $\theta = 0.25$. A framework map of 119 loci or loci clusters (indicated in bold letters along the linkage groups) that could be ordered with a likelihood support ≥1000:1 was assembled covering 95.2% of the estimated genome size. RAPD marker loci that could not be ordered with equal confidence were designated as accessory markers and are listed on the right of the linkage groups along with the approximate cM distance to the closest framework locus. RAPD marker loci are identified by the operon primer code, fragment size in base pairs, fragment amplification intensity and linkage phase (+ or −) (see MATERIALS AND METHODS). A RAPD marker locus showing significant distortion from 1:1 segregation ratio, is indicated by an asterisk.

adjacent markers. A final framework order was accepted when the log-likelihood difference between the initial order and all the alternative local permutations was at least −3.0. Therefore, the framework map orders presented are approximately 1000 times more likely than the next best orders. When local differences of less than −3.0 were observed additional markers were excluded from the framework at the regions of ambiguous ordering, and the analyses performed again. Generally after one or two such iterations, a final order was attained. This final framework order was then compared to the order obtained by simulated annealing implemented by GMENDEL. With the exception of a few two-marker order permutations the framework orders obtained were the same. Segments of the framework where order ambiguities still persisted, were analyzed using error detection functions of MAPMAKER 3.0 that display potential genotyping errors. Gel photos were re-checked for potential scoring errors. In the few instances where the data point was dubious, it was treated as missing data and the ordering analysis performed again.

Markers that could not be placed on the map with a 1000:1 odds were designated as accessory markers and were positioned on the map in relation to the closest framework marker. Their most likely position was obtained by looking for the framework locus that displayed the highest LOD score and lowest two-point θ, or alternatively by looking for the interval with log-likelihood closest to zero. Accessory markers are listed on the right of linkage groups (Figures 3 and 4).

Clustering of markers seems to occur throughout both linkage maps, particularly in *E. urophylla* that displays large clusters on groups 2, 5 and 6. However no formal test for clustering was carried out. Clustering is a common occurrence and has been reported in essentially all relatively dense linkage maps constructed to date irrespective of the organism or technique used to assay DNA polymorphisms: RFLP in tomato (TANKSLEY *et al.* 1992) and common bean (VALLEJOS *et al.* 1992); RAPD in Arabidopsis (REITER *et al.* 1992) or microsatellites in humans (WEISSENBACH *et al.* 1992). Clustering of markers could be an artifact resulting from the limited resolution of our mapping population. For example, while the Arabidopsis map by REITER *et al.* (1992) displayed clustering, the integrated map based on larger populations reported by HAUGE *et al.* (1993) did not. Alternatively, clustering might have a biological basis reflecting suppressed genetic recombination in heterochromatin around the centromeres and/or in telomeric regions as discussed by TANKSLEY *et al.* (1992) following studies on the correlation of genetic and physical structure in the tomato genome (SEGAL *et al.* 1992). As Arabidopsis has a comparatively much smaller proportion of repetitive DNA than Eucalyptus, tomato and humans, less clustering may be expected *a priori* if clustering is due to restricted recombination in regions rich in repetitive DNA.

The great majority of the accessory markers are within 5 cM of the nearest framework marker. It is likely that their ordering ambiguity results largely from the relatively small recombinational distance estimated from a limited number of meioses analyzed. With 62 meioses, the standard error on a recombination fraction of 0.05 is approximately 0.03. However we also observed that 25% of the accessory markers (22 in 88 for *E. grandis*) and 18.5% (25 of 135 in *E. urophylla*) were at distances greater than 6 cM ($\theta \sim 0.05$) and could still not be placed in the framework. The ambiguity in the placement of these markers might be the result of missing data or to errors in genotyping. The overall fraction of missing data points including both data sets was 8.3%. For the framework markers only, this fraction was 4.5%. In our experimental conditions, a genotyping error rate ≤3% was estimated, varying with the RAPD fragment amplification intensity.

Linkage map statistics, estimates of genome size and coverage: Approximately 59% of the markers for *E. grandis* could be placed on the framework defining a total of 142 loci or loci clusters and 1551 cM of total map distance. For *E. urophylla* 47% of the markers could be placed on a framework of 119 loci or loci clusters covering 1101 cM (Figures 3 and 4). Linkage groups were numbered sequentially from the longest to the shortest. For *E. grandis* the average size of linkage groups was 110 ± 35 cM and the range from 41.6 to 156.9 cM. For *E. urophylla* the average size was 99 ± 32 cM and the range from 46.7 to 141 cM. The total number of markers per linkage group (framework and accessory) varied from 6 (group 14) to 30 (group 5) for *E. grandis*, and from 6 (group 11) to 39 (group 6) for *E. urophylla*. The average distance between two framework markers was 12.2 ± 6.3 cM for *E. grandis* and 10.2 ± 6.6 cM for *E. urophylla*. Both maps have a density of 27 cM, which corresponds approximately to a recombination fraction of 0.25 *i.e.*, no interval between two markers is greater than 27 cM.

Total genome size was estimated for both parents using the method of HULBERT *et al.* (1987), as described by VALLEJOS *et al.* (1992). Only framework markers were used in this procedure to avoid an overestimate of genome coverage. For *E. grandis*, the maternal parent, a total map distance of 1620 cM was estimated, of which 1552, *i.e.*, 95.8% were covered by the framework map. For *E. urophylla*, the paternal parent, the total map distance estimated was 1156 cM, of which 1101 cM, *i.e.*, 95.2%, were covered. A reasonably equivalent genome coverage in both species and sexes was therefore achieved with the pseudo-testcross mapping strategy. Given the estimated total map distances and genome sizes of 641 and 646 Mbp/1C (GRATTAPAGLIA and BRADSHAW 1994), the average physical equivalent of 1 cM

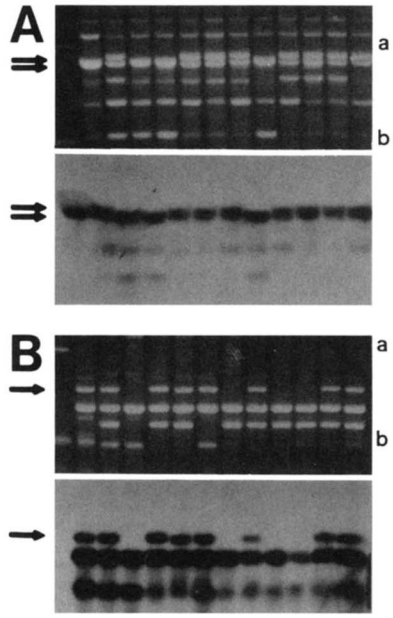


FIGURE 5.—Sequence homology tests of RAPD markers. Panels show the RAPD gel profiles and the corresponding autoradiograms where the indicated RAPD fragments were used as hybridization probes under high stringency to confirm RAPD fragment homology within the parents/progeny set. RAPD assays were performed with genomic DNA of the parents *E. grandis* clone 44 and *E. urophylla* clone 28 (respectively second and third lanes of each panel) and a progeny sample of 10 F_1 individuals. Segregating informative RAPD markers are indicated by arrows. First lane of gel profiles are 1-kb DNA ladder size standards (BRL) (a = 1018 bp; b = 506 bp). (A) Confirmation of allelism between RAPD fragments G14_917, G14_927 and G14_960 amplified in the two parents (see RESULTS for details). (B) Confirmation of homology of a RAPD marker locus (G2_720/3) heterozygous in both parents and therefore segregating 3:1 in the F_1 progeny.

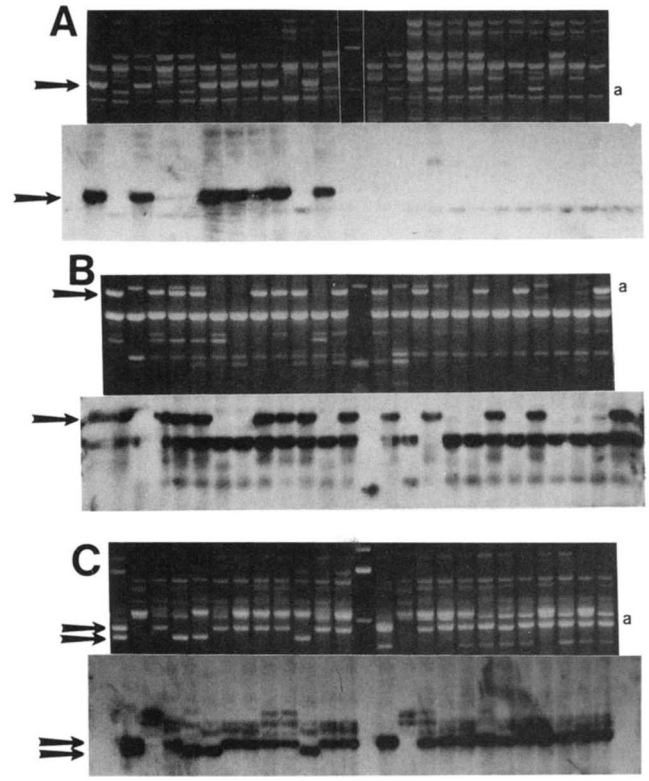


FIGURE 6.—Survey of the occurrence of RAPD markers in different individuals of *E. grandis*. Panels show the RAPD gel profiles and the corresponding autoradiograms where the indicated RAPD fragments were used as hybridization probes under high stringency to confirm RAPD fragment homology. RAPD assay was performed on genomic DNA of two sets of parents and 10 F1 progeny each. Lanes from left to right: *E. grandis* clone 44, *E. urophylla* clone 28; 10 F_1 progeny; 1-kb DNA ladder size standard (BRL) (letter a on the far left indicates the position of the 1018-bp standard), *E. grandis* clone 816/2, *E. urophylla* clone 28, 10 F_1 progeny. (A) RAPD marker N15_1079 present in clone 44 and not in clone 816/2. (B) RAPD marker M5_961 present in both *E. grandis* clones in a heterozygous state thus segregating 1:1 in both crosses. (C) Confirmation of codominance of marker pair A11_980/A11_920. The hybridization experiment also shows that this RAPD marker is also present in clone 816/2 in a heterozygous state with the 980-bp allele and a third size-variant allele, A11_970. This marker locus therefore segregates in both crosses in a codominant fashion, with one allele in common between the crosses.

would correspond to 395 and 559 kbp, respectively, for *E. grandis* and *E. urophylla*.

Confirmation of inheritance and segregation of RAPD markers: The inheritance of segregating RAPD markers from both parents following the pseudo-testcross configuration was confirmed by DNA hybridization experiments (Figures 5 and 6). A case of allelism between RAPD fragments in the two parents was tested and confirmed using DNA hybridization. In *E. grandis*, the RAPD marker G14_927/3 is present in one allelic form with a fragment size of 927 bp. In *E. urophylla*, the same marker is present in two allelic forms: a 917-bp and a 960-bp fragment. In the F_1 a 1:1 segregation is observed for the heterozygote 917/927 (lanes 4, 5 and 9) bp *vs.* the heterozygote 960/927 bp (lanes 6–8 and 10–13). The hypothesized allelism between the three bands in the two parents, was confirmed by probing the RAPD gel blot with the 927-bp RAPD fragment from *E. grandis*, and detecting signal in all three bands (Figure 5, panel A). The same principle was used to confirm the homology of markers that segregated 3:1. Both parents showed bands of equivalent size, and the progeny showed a segregation ratio that fit a 3:1. DNA hybridization of the RAPD gel blot with a fragment from one of the parents confirmed the hypothesis (Figure 5, panel B). A total of 11 markers were found in this configuration, *i.e.*, present in both parents in a heterozygous state therefore segregating 3:1 in the progeny. In principle, such markers could be helpful to define homologies between linkage groups in the two maps. We attempted to do this using GMENDEL that analyzes mixtures of segregation ratios (1:1 and 3:1). Although a few linkages were found at relaxed thresholds, no reliable map position could be established with our sample size. We estimated that the mean amount of information (ALLARD 1956) supplied by

a single individual for such mixture of mating configurations is only ¼ (at $\theta = 0.05$) and 1/3 ($\theta = 0.25$) of the information in the backcross. Therefore to achieve adequate power to estimate linkage in this case, a larger sample size is necessary. Fully informative codominant markers such as microsatellites, RFLPs or isozymes would be highly desirable for this task.

Codominant (size-variant) RAPD markers: Although rare, codominant (size-variant) RAPD markers were found on both maps. Codominant RAPD markers can result from small insertions or deletions between priming sites (WILLIAMS *et al.* 1992). Codominant RAPDs were initially hypothesized from the following observation: (1) both allelic fragments are present in the same parent and are amplified with the same primer; (2) F_1 individuals receive either one or the other allele, *i.e.*, the two RAPD fragments are in repulsion, and (3) no recombinant genotypes are observed in the F_1, *i.e.*, no individuals with both fragments or null for both fragments. In *E. grandis* four marker pairs fit these observations: A10_635/562 (group 5); A11_980/920 (group 7); Y17_525/515 (group 8); Y15_760/740 (group 11). In *E. urophylla* also four pairs were found : U7_1100/850 (group 3); Y20_400/390 (group 5) U13_350/320 (group 5); Z11_550/480 (group 6). DNA hybridization experiments confirmed the codominance of such sets of markers (Figure 6, panel C). Several pairs of RAPD markers were observed that satisfied all the observations outlined above except that the two fragments were amplified with different primers (*e.g.*, in *E. grandis* the pair K9_884/3 and K19_448/3 on group 4). For this category of tightly linked markers in repulsion no DNA hybridization experiments were carried out. Functionally, as pointed out originally by WILLIAMS *et al.* (1990) such pairs of markers could also be used as a single codominant marker. Overall, however, the frequency of truly codominant and functionally codominant RAPD markers for the *Eucalyptus* species surveyed remains below 3%.

Survey of the presence and allelic state of RAPD markers in different individuals: A subset of markers from the *E. grandis* clone 44 map were surveyed for their presence and allelic state in a second individual tree of the same species (*E. grandis* 816/2) by analyzing their segregation in a second F_1 progeny set involving 816/2 as a parent. From a total of 112 RAPD markers surveyed, 37 (33%) were found to amplify in the second tree. Of these, 20 were also in a heterozygous state and segregated 1:1 in the F_1, while 17 were homozygous, *i.e.*, did not segregate. DNA hybridization experiments were carried out for 5 markers that were not shared (Figure 6, panel A), and 10 that were shared confirming the homology (Figure 6, panels B and C). Only one marker of the ones tested was found to be misinterpreted, *i.e.*, the RAPD bands were scored as being the same but in fact were not homologous. This was a relatively large (1500 bp) fragment. Misinterpretations of this kind are more likely for larger fragments that are not as efficiently size fractionated as smaller fragments. In carrying out this kind of RAPD marker survey we found it imperative to run the gels until the dye marker has run for at least 15 cm, to minimize errors resulting from co-migrating fragments.

For *E. urophylla*, after surveying 34 randomly chosen RAPD markers, it was found that all of them were present in what was thought to be a different individual of the same species. Moreover, all the markers surveyed also segregated in the second F_1 progeny. The possibility that the same *E. urophylla* clone 28 had actually been used as the male parent in the second cross was tested and confirmed. A subset of five markers that were not recombinationally separated in a locus cluster on group 5 (defined by marker U13_350/2) and a second subset of four markers in a locus cluster on group 8 (defined by marker M4_477/2) were surveyed for linkage on a set of 16 progenies. All the markers were found to be present and no recombinants were found. Further DNA hybridization experiments also confirmed that in fact the same *E. urophylla* (clone 28) had been used as the male parent in both crosses (data not shown). This result did not allow us to explore the extent of conservation of RAPD markers in different individuals of *E. urophylla*, however it was useful to confirm the stable behavior of RAPD markers in terms of segregation and linkage relationships in a second cross involving the same individual tree.

Characterization of genomic sequence complexity of RAPD marker loci: Over 50% of the 48 RAPD fragments surveyed were found to amplify from low copy genomic regions (1–10 copies) and less than 10% originated from very highly repeated regions (≤1000 copies). Approximately equal frequencies (~20%) were found for fragments amplified from moderately repeated (10–100) and highly repeated regions (100 to 1000) (Figure 7, panels A and B). Similar estimates of genomic sequence complexity of RAPD marker loci was observed in soybean (WILLIAMS *et al.* 1990) and Arabidopsis (REITER *et al.* 1992). Based on 48 data points, a simple correlation analysis was carried out between the following variables: RAPD fragment size in base pairs, amplification intensity score and copy number class. The results were as follows: fragment size × intensity score $r = 0.25$; fragment size × copy number $r = 0.04$; intensity score × copy number $r = 0.18$. In conclusion, no significant correlation ($\alpha = 0.05$) was found for any of the three pairwise analyses, suggesting no particular dependency of the fragment size or amplification efficiency of RAPD marker loci on the complexity of the genomic region sampled.

It is important to point out that the sample of RAPD fragments surveyed in this experiment are not randomly chosen RAPD fragments, rather they correspond to a

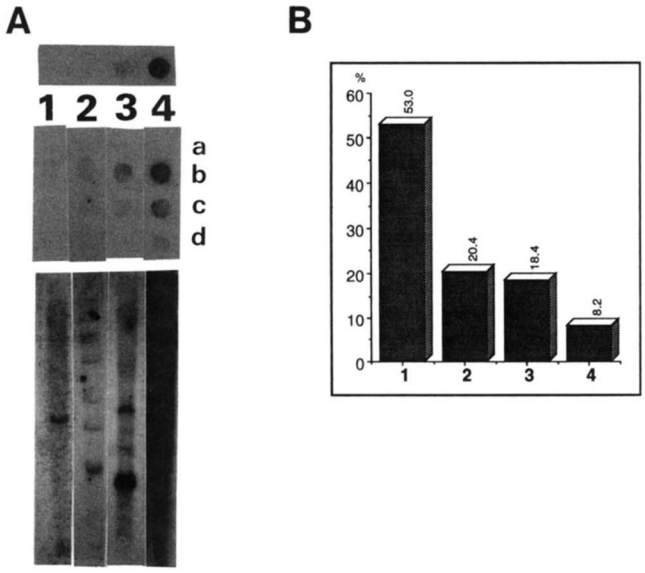


FIGURE 7.—Characterization of genomic sequence complexity of mapped RAPD marker loci. (A) Top horizontal dot blot corresponds to a control blot with a reconstruction experiment where signals indicate a sequence represented by (1) 1 copy; (2) 10 copies; (3) 100 copies; (4) 1000 copies (see MATERIALS AND METHODS for details). Vertical dot blots and corresponding *Eco*RI Southern blots show RAPD fragments classified in the following copy number classes: (1) low copy (1–10 copies); (2) moderately repetitive (10–100 copies); (3) highly repetitive (100–1000 copies) and (4) very highly repetitive (>1000 copies). Dots contain (a) negative control, 5 μg herring sperm DNA; (b) 5 μg Eucalyptus genomic DNA; (c) 0.5 μg; (d) 0.05 μg. (B) Frequency histogram of genome complexity classes of mapped RAPD markers based on a sample of 48 RAPD markers. Numbers on *x*-axis correspond to the same copy number classes described above.

subset of fragments that behave as genetic markers, map to a single location and therefore originate from a unique site in the genome. WOLFF *et al.* (1993) found that 48% of random *Pst*I genomic clones of *E. grandis* and *E. urophylla* were useful as RFLP probes. In our study we found that 53% of the mapped RAPD fragments amplified from low copy regions and could potentially be used as RFLP probes. These results indicate that the genomic library of mapped RAPD fragments, obtained in Eucalyptus as a byproduct of this mapping experiment, closely resembles a genomic library of RFLP probes constructed by the traditional approach.

DISCUSSION

Pseudo-testcross mapping strategy using RAPD markers: We have used a "pseudo-testcross" mapping strategy in combination with the RAPD assay to construct the first reported linkage maps for species of Eucalyptus. In a cross between heterozygous parents, many single-dose polymorphic markers will be heterozygous in one parent, null in the other and therefore segregate 1:1 in their progeny as in a testcross. We use the name "pseudo-testcross" for this strategy because the testcross mating configuration of the markers is not known *a priori* as in a conventional testcross where the tester is homozygous recessive for the locus of interest. Rather, the configuration is inferred *a posteriori* after analyzing the parental origin and genetic segregation of the marker in the progeny of a cross between highly heterozygous parents with no prior genetic information. When this inference is done for both parents involved in the cross, the term "two-way pseudo-testcross" is more appropriately used.

RITTER *et al.* (1990) described the theoretical background for linkage analysis of markers segregating in crosses between heterozygous parents. As mentioned in that work, map construction in allogamous plant species for which only heterozygous individuals are available can make use of single-dose polymorphic markers behaving as dominant markers in an F_1, segregating 1:1 for the presence or absence of the fragment. These markers were used for genetic mapping in potato (BONIERBALE *et al.* 1988; GEBHARDT *et al.* 1989), and recently allowed genetic mapping in polyploid sugar cane (WU *et al.* 1992; DASILVA *et al.* 1993; SOBRAL and HONEYCUTT 1993; AL-JANABI *et al.* 1993). We and others observed this same mating configuration when analyzing genetic segregation of RAPD markers in F_1 crosses of forest and fruit trees, and suggested its wide applicability for genetic mapping in this group of highly heterozygous largely undomesticated species (CARLSON *et al.* 1991; GRATTAPAGLIA *et al.* 1992a; ROY *et al.* 1992; LAWSON *et al.* 1992).

The pseudo-testcross mapping strategy is conceptually simple to implement and can be applied with any type of molecular marker. However its potential can be better explored with the efficiency of the RAPD assay in pre-screening marker polymophisms in search of the informative testcross configurations. The fact that the RAPD assay is sensitive to single base changes, contributes to a higher efficiency in scanning the genomes for polymorphisms. Moreover, the fact that RAPD detects only one allele at a locus facilitates the occurrence of pseudo-testcross configurations, because the necessary null genotype of one of the parents actually corresponds to undetected alleles. In addition to that, the RAPD assay is technically simple and fast to perform facilitating the initial screening step. Following our screening procedure in Eucalyptus, 36 arbitrary primers could be easily screened in a single working day, yielding an estimated 1.82 markers/primer, *i.e.*, 65 markers from both parents taken together. Finally, the segregation ratio observed for a dominant RAPD marker in this configuration has the same information content as that of a codominant marker. Evidently, a highly polymorphic, multiallelic marker that detected all four allelic variants of the mating configuration (*e.g.*, sequence tagged microsatellite site), would contain more genetic information (MORGANTE and OLIVIERI 1993).

The use of RAPD markers in a pseudo-testcross configuration is a general strategy for the construction of genetic linkage maps in outbred forest trees as well as in

any highly heterozygous living organisms. It can be immediately applied to any species without any prior genetic information. The only requirements are sexual reproduction between two individuals that results in the generation of a progeny large enough to allow the estimation of recombination frequencies between segregating markers. Its efficiency will be directly proportional to the level of genetic heterozygosity of the species under study, which is a function of the mating system, and the genetic divergence between the individuals crossed. In our study we employed an interspecific cross between highly heterozygous individuals from two closely related outcrossing species, thus increasing the probability of finding pseudo-testcross marker configurations. We found only 11 markers heterozygous in both parents thus segregating 3:1 compared to 558 markers in a testcross configuration, segregating 1:1.

The pseudo-testcross strategy should also be efficient at the intraspecific level and increasingly so with crosses of genetically divergent individuals from geographically distinct origins. In a survey of 112 mapped markers, we found that only 33% were shared between two individuals of the same species and different provenances. We suggest that at the intraspecific level, the mapping efficiency of the pseudo-testcross strategy, measured by the number of informative markers/arbitrary primer should reach between 60 and 70% of the one reported in this study, that is, ~2.4 instead of 3.69 markers/selected primer. With individuals from the same population, this number will tend to be lower, as more markers will be shared. In a group of 38 heterozygous clones of *Solanum tuberosum* the informativeness of RFLP probes for direct segregation analysis in F_1 populations varied from 49 to 95%, indicating that linkage mapping using F_1 progeny should be feasible for most combinations (GEBHARDT *et al.* 1989). Testcross RAPD marker configurations were often observed at the intraspecific level in other highly heterozygous forest tree species, however, no estimates of frequencies per arbitrary primer were given (CARLSON *et al.* 1991; ROY *et al.* 1992).

The pseudo-testcross strategy basically extends the haploid mapping approach used for conifers, to any other angiosperm tree species. The final result is essentially the same, *i.e.*, linkage maps for individual trees; however, it requires performing a controlled cross. On the other hand it is more time and cost efficient since gametic segregation from two individuals or twice the heterozygosity is surveyed simultaneously in the same PCR reaction, both in the primer screening and in the mapping phase. Therefore, even in conifers, the pseudo-testcross could potentially be the mapping strategy of choice for quickly generating single-tree linkage maps.

Genetic linkage maps of single individuals: The genetic linkage maps constructed in this study (Figures 3 and 4) are individual-specific. The pseudo-testcross strategy is specifically based on the selection of single-dose markers present in one parent and absent in the other. In our maps, no RAPD markers are in common and so it is not possible to determine homologies of linkage groups in the two maps or integrate the two maps into one. Overlap of RAPD marker occurrence and linkage relationships in genetic maps of different individuals will depend on the presence of the same RAPD marker loci and their allelic state. While at the interspecific level, the overlap will be very low, at the intraspecific level, it will be increasingly high as individuals from the same population are used. In this study we found that 33% of the mapped markers in *E. grandis* were present in a second individual of the same species but from a very distinct origin, and 54% of those were also in a heterozygous state. Indirect evidence for the occurrence of the same RAPD markers across different individuals of the same population come also from studies that employed RAPD markers to estimate outcrossing rates in stands of *E. urophylla* (GRATTAPAGLIA *et al.* 1992a) and *Datisca glomerata* (FRITSCH and RIESEBERG 1992).

To integrate linkage maps constructed by the pseudo-testcross strategy, multiallelic codominant markers with alleles segregating from both parents would be most efficient, providing a set of common loci which could be used as locus bridges. In this study, 11 dominant RAPD markers were shared by the parents and could in principle be used to assign linkage groups homologies between the two maps. However, due to the low information content of the mixed mating configuration, a larger sample size would be necessary to achieve adequate power to accomplish such assignment. More than 50% of the RAPD fragments mapped in this study are low copy and could potentially be used to detect codominant RFLPs. Although such markers could be helpful in connecting linkage groups, a complete map merging would still be fairly difficult to achieve, since correct locus ordering among the markers not in common between homologous linkage groups would not be known (B. H. LIU, personal communication). Statistical integration of mapping data from different populations of Arabidopsis was shown to be problematic, especially in regions containing low densities of common markers between maps (HAUGE *et al.* 1993). As mentioned earlier, multiallelic markers such as microsatellites would be very powerful and desirable for this purpose. In the context of the breeding applications envisaged for these linkage maps in forest tree breeding, map merging is not immediately necessary, but it will become important in subsequent generations of breeding and selection (see discussion below).

No morphological traits or other single gene traits that could potentially be placed on these maps, segregated in the population used for map construction. To our knowledge no simply inherited traits are presently

known in Eucalyptus that could be placed on any genetic map, with the exception of isozyme loci. The same screening procedure with parents and a subset of progeny could be used for mapping isozymes loci. Informative configurations of isozymes genotypes in the parents would result in either 1:1 or 1:2:1 segregation ratios in the F_1, and mapping of the isozyme locus could be achieved in only one or both maps respectively. The existing RAPD linkage maps provide a scaffold where even distant linkage relationships of isozymes could be determined. On the other hand, isozyme loci would provide anchor loci for single-tree map comparisons and merging.

We obtained equivalent genome coverages on the two maps in spite of different estimates of total map distances. This suggests that the difference in total map distance observed between the two species and sexes are biologically significant. However no distinction is possible at this point between a species specific, sex specific or individual specific difference in genetic recombination. If applied within species, however, the pseudo-testcross strategy should provide a valuable tool to study specific differences in general recombination rate. Genome sizes estimated in our study are well within the range of several other species (listed by NODARI *et al.* 1993). However, genome coverages found in our study are slightly higher than those found in other maps. A good comparative example in this respect is common bean, that has genome characteristics similar to eucalypts ($n = 11$ chromosomes; genome size around 600 Mbp). Approximately 80% of the genome could be covered with a framework map of 145 RFLP loci (VALLEJOSAMA *et al.* 1992). Besides intrinsic biological differences in levels of DNA polymorphism and rates of recombination, one of the possible reasons for the observed difference in genome coverage could be the result of a more efficient genome sampling by RAPD markers as compared to the RFLP technique, particularly for genomic regions rich in repetitive DNA.

Framework maps at a likelihood ratio support $\geq$1000:1 were constructed (Figures 3 and 4). This presentation of the data is convenient for selecting a subset of evenly spaced framework markers to initially scan the genome for quantitative trait locus (QTL) mapping. Recent simulation studies have shown that wide marker spacings of 20 or even 50 cM are optimal for this task (DARVASI *et al.* 1993). A more focused search for the exact position of QTLs can then be done with the available nearby accessory markers. If more markers are still needed in a region of interest, genetic walking based on genotype pooling techniques could be used (MICHELMORE *et al.* 1991; GIOVANNONI *et al.* 1991; REITER *et al.* 1992).

Toward marker assisted breeding strategies in forest trees: It has been long recognized that one of the problems facing marker assisted breeding in outbred species such as forest trees is the linkage equilibrium between marker loci and genetic loci of interest (SOLLER 1978; BECKMANN and SOLLER 1983; NEALE and WILLIAMS 1991; LANDE and THOMPSON 1990; STRAUSS *et al.* 1992). With linkage equilibrium, marker-trait associations established in one cross, would not hold in a second pedigree, since marker and QTL alleles would be randomly associated. As pointed out earlier, one solution would be to construct maps for each genotype in the breeding population (NEALE and WILLIAMS 1991; GRATTAPAGLIA *et al.* 1992b). This was considered a task that depended on significant advances in the ability to obtain marker data. The RAPD technology provided this advance by allowing the construction of a 200 marker linkage map in *Pinus taeda* in 6 person-months of work (GRATTAPAGLIA *et al.* 1991). In this study, with the pseudo-testcross strategy and RAPD markers we constructed two linkage maps simultaneously in approximately 5 person-months of work.

The ability to construct genetic linkage maps quickly in any forest tree opens the way to the heterodox proposal of constructing maps for individual trees in a breeding population. The paradigm of an index linkage map for a species is an attractive one for comparative mapping applications. However it does not seem adequate as the initial approach for establishing marker/traits associations for breeding in allogamous populations with a wide genetic base such as those of forest trees. Rather, the progressive accumulation of individual linkage maps with subsets of common markers among them will make obvious the relationships of linkage groups in different maps. This will eventually lead to a unified map where general regions associated with trait expression could be identified. Multiallelic codominant markers such as microsatellites would then be highly desirable specifically bracketing such regions to facilitate their manipulation in breeding. Although such "population level" or general QTLs should exist, their relative importance in the overall level of genetic variation in quantitative traits in forest trees is still unclear. The identification and manipulation of QTLs specific to individual trees might emerge as being more important for the advancement of quantitative traits by marker assisted breeding.

Based on the proposal of individual-specific linkage maps, the integration of mapping information into tree breeding programs would involve four basic steps, briefly: (1) construction of moderate density individual tree maps for elite genotypes, in a two-by-two fashion using the "two-way pseudo-testcross" strategy; (2) localization of favorable alleles at qualitative and quantitative trait loci of interest on these maps, by analyzing the performance of an extended set of the full-sib family used for map construction. This extended set of progeny would be initially genotyped only for a subset of evenly spaced framework markers, followed by a finer search

with accessory markers in regions of interest. Alternatively half-sib families of the mapped individuals could also be used for this purpose. In this step, retrospective QTL analysis using existing full and half-sib families at harvest age would be highly preferable in order to gather the necessary quantitative data in acceptable time. Power and precision in QTL mapping would be greatly enhanced by using large family block plantations (>1000 individuals) and clonal replication of genotypes when possible; (3) validation of marker-trait associations by replication and prediction experiments; (4) marker assisted selection of progeny, or retrospective selection of parents for planned recombination in subsequent generations of breeding. Given that close linkages are established in the mapping phase, the decay of marker-trait associations with time, would not be of immediate concern in the context of the long generations of tree breeding.

We wish to thank the members of the Forest Biotechnology Group (NCSU) in particular José X. Chaparro and David O'Malley for the animated and valuable discussions, Ross Whetten for the advice with the characterization of RAPD fragments and Ben H. Liu for the statistical counseling. We also want to acknowledge Bruce Weir for the encouragment throughout this work and critical review of the manuscript and Scott Tingey for providing the latest Mac versions of MAPMAKER. D.G. has been supported by a doctoral fellowship from CAPES (Brazilian Ministry of Education) and a research grant from Aracruz Florestal S.A. We thank the NCSU Forest Biotechnology Industrial Associates Consortium and the U.S. Department of Agriculture Competitive Grants Program for their support. This work was partially supported by U.S. Department of Agriculture award 91-37300-6431.

LITERATURE CITED

Al-Janabi, S. M., R. J. Honeycutt, M. McClelland and B. W. S. Sobral, 1993 A genetic linkage map of *Saccharum spontaneum* L. 'SES 208.' Genetics **134:** 1249–1260.

Allard, R., 1956 Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24:** 235–278.

Beckmann, J. S., and M. Soller, 1983 Restriction fragment length polymorphisms in genetic improvement methodologies, mapping and costs. Theor. Appl. Genet. **67:** 35–43.

Boland, D. J., M. I. H. Brooker, G. M. Chippendale, N. Hall, B. P. M. Hyland *et al.*, 1984 *Forest trees of Australia*, Ed. 4. Nelson-CSIRO, Melbourne.

Bonierbale, M. W., R. L. Plaisted and S. D. Tanksley, 1988 RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120:** 1095–1103.

Bradshaw, H. D., and R. F. Stettler, 1993 Molecular genetics of growth and development in *Populus*. I. Triploidy in hybrid poplars. Theor. Appl. Genet. **86:** 301–307.

Caetano-Anollés, G., B. J. Bassam and P. M. Gresshoff, 1991 High resolution DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. Bio/Technology **9:** 553–557.

Carlson, J. E., L. K. Tulsieram, J. C. Galubitz, V. W. K. Luk, C. Kauffeldt *et al.*, 1991 Segregation of random amplified DNA markers in F_1 progeny of conifers. Theor. Appl. Genet. **83:** 194–200.

Chaparro, J. X., D. J. Werner, D. M. O'Malley and R. R. Sederoff, 1994 Targeted mapping and linkage analysis of morphological, isozyme and RAPD markers in peach. Theor. Appl. Genet. **87:** 805–815.

Conkle, M. T. 1981 Isozyme variation and linkage in six conifer species, pp. 11–17 in *Proceedings of the Symposium on Isozymes in North American Forest Trees and Forest Insects*. Pacific Southwest Experiment Station, USDA Genetics Technical Report PSW-48.

Darvasi, A., A. Weinreb, V. Minke, J. I. Weller and M. Soller, 1993 Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics **134:** 943–951.

DaSilva, J. A. G., M. E. Sorrels, W. L. Burnquist and S. D. Tanksley, 1993 RFLP linkage map and genome analysis of *Saccharum spontaneum*. Genome **36:** 782–791.

Devey, M. E., K. D. Jermstad, C. G. Tauer and D. B. Neale, 1991 Inheritance of RFLP loci in a loblolly pine three-generation pedigree. Theor. Appl. Genet. **83:** 238–242.

Doyle, J. J., and J. L. Doyle, 1987 Isolation of plant DNA from fresh tissue. Focus **12:** 13–15.

Eldridge, K., J. Davidson, C. Harwood and G. Van Wyk, 1993 *Eucalypt Domestication and Breeding*. Oxford University Press, Oxford.

Fripp, Y. J., A. R. Griffin and G. F. Moran, 1987 Variation in allele frequencies in the outcross pollen pool of *Eucalyptus regnans* F Muell. throughout a flowering season. Heredity **59:** 161–171.

Fritsch, P., and L. H. Rieseberg, 1992 High outcrossing rates maintain male and hermaphrodite individuals in populations of the flowering plant *Datisca glomerata* Nature **359:** 633–636.

Gebhardt, C., E. Ritter, T. Debener, U. Schachtschabel, B. Walkemeier *et al.*, 1989 RFLP analysis and linkage mapping in *Solanum tuberosum*. Theor. Appl. Genet. **78:** 65–75.

Giovannoni, J. J., R. A. Wing, M. W. Ganal and S. D. Tanksley, 1991 Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. Nucleic Acids Res. **19:** 6553–6558.

Grattapaglia, D., 1994 Applications of molecular markers in plant breeding, in *Genetic Mapping of Quantitatively Inherited and Economically Important Traits in Eucalyptus*. Ph.D. Dissertation, North Carolina State University, Raleigh.

Grattapaglia, D., and H. D. Bradshaw, Jr., 1994 Nuclear DNA content of commercially important *Eucalyptus* species and hybrids. Can. J. For. Res. (in press).

Grattapaglia, D., and R. R. Sederoff, 1992 Pseudo-testcross mapping strategy in forest trees: single tree RAPD maps of *Eucalyptus grandis* and *E. urophylla*. Plant Genome I, San Diego, Calif., poster #47.

Grattapaglia, D., P. Wilcox, J. X. Chaparro, D. O'Malley, S. McCord *et al.*, 1991 A RAPD map of loblolly pine in 60 days. Third International Congress of the International Society for Plant Molecular Biology, abstract 2224.

Grattapaglia, D., D. M. O'Malley and R. R. Sederoff, 1992a Multiple applications of RAPD markers to genetic analysis of *Eucalyptus* sp., pp. 436–450 in *Proceedings of IUFRO International Conference "Breeding Tropical Trees,"* Section 2.02–08, Cali, Colombia.

Grattapaglia, D., J. X. Chaparro, P. Wilcox, S. McCord, D. Werner *et al.*, 1992b Mapping in woody plants with RAPD markers: applications to breeding in forestry and horticulture, pp. 37–40 in *Proceedings of the Symposium on Applications of RAPD Technology to Plant Breeding*. CSSA/ASHS/AGA.

Hauge, B. M., S. M. Hanley, S. Cartinhour, J. M. Cherry, H. M. Goodman *et al.*, 1993 An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J. **3:** 745–754.

Hulbert, S. H., T. W. Ilot, E. J. Legg, S. E. Lincoln, E. S. Lander *et al.*, 1987 Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics **120:** 947–858.

Kazan, K., J. M. Manners and D. F. Cameron, 1993 Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. Genome **36:** 50–56.

Kesseli, R. V., I. Paran and R. W. Michelmore, 1990 Genetic linkage map of lettuce (*Lactuca sativa*, $2n = 18$), pp. 6100–6102 in *Genetic Maps*, edited by S. J. O'Brien. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Lande, R., and R. Thompson, 1990 Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics **124:** 743–756.

Lander, E. S., P. Green, J. Abrahamson, A. Baarlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Lawson, D., M. Hemmat and N. Weeden, 1992 Identification of genes influencing morphological and physiological traits in apple. Abstracts of the International Conference on the Plant Genome I, San Diego, Calif., poster #148.

LIU, B. H., and S. J. KNAPP, 1990 GMENDEL: a program for Mendelian segregation and linkage analysis of individual or multiple progeny using log-likelihood ratios. J. Hered. **81:** 407.

MICHELMORE, R. W., I. PARAN and R. V. KESSELI, 1991 Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA **88:** 9828–9832.

MORAN, G. F., and J. C. BELL, 1983 *Eucalyptus*, pp. 423–441 in *Isozymes in Plant Genetics and Breeding*, Part B, edited by S. D. TANKSLEY and T. J. ORTON. Elsevier, Amsterdam.

MORAN, G. F., J. C. BELL, M. BYRNE, M. E. DEVEY and D. SMITH, 1992 Molecular marker maps in *Pinus radiata* and *Eucalyptus nitens* for application in tree breeding programs. Plant Genome I, San Diego, Calif., poster #93.

MORGANTE, M., and A. M. OLIVIERI, 1993 PCR-amplified microsatellites as markers in plant genetics. Plant J. **3:** 175–182.

MURASHIGE, T., and F. SKOOG, 1962 A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. **15:** 473–497.

NEALE, D. B., and C. G. WILLIAMS, 1991 Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. Can. J. For. Res. **21:** 545–554.

NELSON, C. D., W. L. NANCE and R. L. DOUDRICK, 1993 A partial genetic linkage map of slash pine (*Pinus elliotti* Engelm var. *elliottii*) based on random amplified polymorphic DNAs. Theor. Appl. Genet. **87:** 145–151.

NODARI, R. O., S. M. TSAI, R. L. GILBERTSON and P. GEPTS, 1993 Towards an integrated linkage map of common bean. II. Development of an RFLP-based linkage map. Theor. Appl. Genet. **85:** 513–520.

REITER, R. S., J. G. K. WILLIAMS, K. A. FELDMANN, J. A. RAFALSKI, S. V. TINGEY *et al.*, 1992 Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc. Natl. Acad. Sci. USA **89:** 1477–1481.

RHOADS, D. D., and D. J. ROUFA, 1990 Seqaid II 3.80. Molecular Genetics Laboratory, Kansas State University, Manhattan.

RITTER, E., C. GEBHARDT and F. SALAMINI, 1990 Estimation of recombination frequencies and construction of linkage maps from crosses between heterozygous parents. Genetics **125:** 645–654.

ROY, A., N. FRASCARIA, J. MACKAY and J. BOUSQUET, 1992 Segregating random amplified polymorphic DNAs (RAPDs) in *Betula alleghaniensis*. Theor. Appl. Genet. **85:** 173–180.

SAMPSON, J. F., S. D. HOPPER and S. H. JAMES, 1990 Temporal variation in allele frequencies in the pollen pool of *Eucalyptus rhodantha*. Heredity **65:** 189–199.

SEGAL, G., M. SARFATTI, M. A. SCHAFFER, N. OVI, D. ZAMIR *et al.*, 1992 Correlation of genetic and physical structure in the region surrounding the *I2 Fusarium oxysporum* resistance locus in tomato. Mol. Gen. Genet. **231:** 173–185.

SOBRAL, B. W. S., and R. J. HONEYCUTT, 1993 High output genetic mapping of polyploids using PCR-generated markers. Theor. Appl. Genet. **86:** 105–112.

SOLLER, M., 1978 The use of loci associated with quantitative traits in dairy cattle improvement. Anim. Prod. **27:** 133–139.

SONG, Y., and C. A. CULLIS, 1992 RFLP and RAPD mapping of *Eucalyptus globulus*. Plant Genome I, San Diego, Calif., poster #133.

SOUTHERN, E. M. 1975 Detection of specific sequences among DNA fragments separated by electrophoresis. J. Mol. Biol. **98:** 503–517.

STEANE, D. A., A. K. WEST, B. M. POTTS, J. R. OVENDEN and J. B. REID, 1991 Restriction fragment length polymorphism in chloroplast DNA from six species of *Eucalyptus*. Aust. J. Bot. **39:** 399–414.

STRAUSS, S. H., R. LANDE and G. NAMKOONG, 1992 Limitations of molecular-marker aided selection in forest tree breeding. Can. J. For. Res. **22:** 1050–1061.

STUBER, C. W., 1992 Biochemical and molecular markers in plant breeding. Plant Breed. Rev. **9:** 37–61.

TANKSLEY, S. D., M. W. GANAL, J. P. PRINCE, M. C. DE VICENTE, M. W. BONIERBALE *et al.*, 1992 High density molecular linkage maps of the tomato and potato genomes. Genetics **132:** 1141–1160.

TINGEY, S. V., and J. P. DELTUFO, 1992 Genetic analysis with random amplified polymorphic DNA markers. Plant Physiol. **101:** 349–352.

TORRES, A. M., N. F. WEEDEN and A. MARTIN, 1993 Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. Theor. Appl. Genet. **85:** 937–945.

TULSIERAM, L. K., J. C. GLAUBITZ, G. KISS and J. CARLSON, 1992 Single tree genetic linkage mapping using haploid DNA from megagametophytes. Bio/Technology **10:** 686–690.

VALLEJOS, C. E., N. S. SAKIYAMA and C. D. CHASE, 1992 A molecular marker-based linkage map of *Phaseolus vulgaris* L. Genetics **131:** 733–740.

WEISSENBACH, J., G. GYPAY, C. DIB, A. VIGNAL, J. MORISSETTE *et al.*, 1992 A second-generation linkage map of the human genome. Nature **359:** 794–801.

WELSH, J., and M. MCCLELLAND, 1990 Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. **18:** 7213–7218.

WILLIAMS, J. G. K., M. K. HANAFEY and J. A. RAFALSKI, 1993 Genetic analysis using randam amplified polymorphic DNA markers. Methods Enzymol. **218:** 704–740.

WILLIAMS, J. G. K., A. R. KUBELIK, K. J. LIVAK, J. A. RAFALSKI and S. V. TINGEY, 1990 DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. **18:** 6531–6535.

WOLFF J. L. C., S. W. MCDOWELL and M. A. MACHADO, 1993 Analysis of the restriction fragment length polymorphism of *E. urophylla*, *E. grandis* and *E. citriodora*. Braz. J. Genet. **16:** 431–439.

WU, K. K., W. BURNQUIST, M. E. SORRELLS, T. L. TEW, P. H. MOORE *et al.*, 1992 The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor. Appl. Genet. **83:** 294–300.

YEH, F. C., A. BRUNE, W. M. CHELIAK and D. C. CHIPMAN, 1983 Mating system of *Eucalyptus citriodora* in a seed production area. Can. J. For. Res. **13:** 1051–1055.

Communicating editor: W. G. HILL